

Supporting Text

The Need for Analysis in Terms of Geometry Rather Than Contrast Alone. As briefly noted in the text, when considering lines it is only natural to think of stimuli generated by luminance contrasts in a scene. Luminance contrast boundaries (i.e., edges) have traditionally been the focus of vision research, primarily because edges typically correspond to object boundaries and are therefore considered “information-rich.” Furthermore, there is much evidence that observers preferentially attend to contrast boundaries in normal viewing, which generally elicit much stronger responses in visual cortical neurons than stimuli that lack contrast boundaries.

Despite the obvious importance of edges in vision, lines in images (straight or otherwise) are not limited, of course, to those that happen to coincide with luminance contrast boundaries. In terms of geometry, a straight line is simply a set of points with positions in space that conform to a linear progression. The points in space that make up such lines are obviously perceived when observers view surfaces in any scene; the diagonals of a square surface, for example, don’t need to be drawn as contrast boundaries on the surface for the observer to perceive these particular linear aspects of the square or any other parallelogram. Thus, with respect to behavior, appreciating lines that do not happen to coincide with object edges is just as important as appreciating those that do, counterintuitive though this may seem. Although lines defined by geometry alone do not stimulate neurons specifically activated by contrast boundaries (“edge detectors”), the perception that follows (and an appreciation of the empirical significance of the diagonals of the square, for instance) is presumably just as important for successful navigation in typical environments as is the subset of lines made explicit by contrast.

An engaging example of the behavioral need to appreciate such lines is provided by the children’s game called “jacks.” For those not familiar with the game, the challenge is to pick up a collection of small objects off a surface and catch a rubber ball before it bounces a second time. Successfully picking up a number of jacks during the brief interval that the ball is in the air (one wins by picking up more jacks than the opponent) demands exquisite visually guided behavior predicated on an appreciation of the spatial arrangement of the jacks scattered across the playing surface (in addition to good motor skills). Children (and willing adults) have no problem understanding the relevant geometrical relationships (e.g., the lengths and orientations of the “lines” between the jacks), although these sets of points on the relevant surface produce no luminance contrast boundaries.

Accordingly, an analysis of scene geometry based only on edges (which are included in our straight-line samples) would be flawed by virtue of having considered only a small subset of the real-world lines pertinent to successful behavior (see ref. 1 for an example of where the edges in scenes were extracted and compared to the complete set of straight lines in the database). We therefore included all geometrically straight lines in the analysis.

Comparison of Empirical Ranking Theory and Bayesian Decision Theory. Readers may also wish to know how the present approach, empirical ranking theory, differs from the far better known framework of Bayesian decision theory. Because the distinctions are especially important in how visual processing is conceived, we include this brief account.

Bayesian Decision Theory. As indicated in the article, the rationale for any statistical approach to vision is the fact that the information in a retinal image cannot specify its real-world sources (the inverse optics problem). The most popular approach to understanding how vision contends with this problem has been Bayesian decision theory, which is predicated on the idea that the visual system makes inferences (usually taken to be unconscious) about the properties of the physical world based on prior experience (2-9). To summarize this approach briefly, Bayes' theorem formally states the logical relationships in making valid inferences and is usually written in the form

$$P(H|E) = \frac{P(H) \cdot P(E|H)}{P(E)}$$

where H is a hypothesis, E is the evidence pertinent to its validity, and P is probability. The first term on the right side of the Bayes' equation, $P(H)$, is the prior probability distribution and is a statistical measure of confidence in the hypothesis, absent any present evidence pertinent to its truth or falsity. With respect to vision, the prior describes the relative probabilities of different physical states of the world pertinent to retinal images, i.e., the relative frequency of occurrence of various object sizes, illuminants, surface reflectance values, and so on. The second term, $P(E|H)$, is called the likelihood function. If hypothesis H were true, this term indicates the probability that the evidence E would have been available to support it. In the context of vision, given a particular state of the physical world (i.e., a particular combination of illumination, reflectance properties, object sizes, etc.), the likelihood function describes the probability that the state would generate the retinal projection in question. The product of the prior and the likelihood function, divided by a normalization constant, $P(E)$, gives the posterior probability distribution, $P(H|E)$. The posterior distribution defines the probability of hypothesis H being true, given the evidence E . In vision, the posterior probability distribution thus indicates the relative probability of a given retinal image having been generated by one or another different physical realities.

To illustrate the application of Bayes' theorem to understanding the perceptions elicited by the simple geometrical stimuli in the present article, consider an angle having a particular projected subtense in the retinal image. In terms of physical measurement, the projected subtense in the image is the result of the physical subtense of the object and its orientation in space. The goal of a Bayesian approach to explaining the apparent subtense experienced by an observer in this example is to generate the pertinent posterior probability distribution $P(\text{physical subtense, orientation}|\text{projected subtense})$. The first step, then, is to compute the prior distribution $P(\text{physical subtense, orientation})$, i.e., the probability distribution of the various conjunctions of physical subtense and orientation

of objects in the physical world. In principle, this distribution could be generated by sampling a large number of instances in typical physical environments, measuring the relevant variables in each instance (a difficult but not impossible task). The next step would be to derive the likelihood function $P(\text{projected subtense}|\text{physical subtense, orientation})$, which describes, for each possible combination of the physical variables, the probability that the combination generated the projected subtense in the image. Finally, the posterior distribution, $P(\text{physical subtense, orientation}|\text{projected subtense})$, is obtained by multiplying the prior distribution by the likelihood function. The posterior distribution therefore describes the relative probabilities of all the possible physical sources that could have generated the specific angle in the image under consideration. In short, a Bayesian framework entails the set of physical source(s) capable of generating a given retinal image and the relative probabilities of their actually having done so; the percepts predicted therefore correspond to “explicit models of world structure” (ref. 6, p. 7).

Because the posterior distribution indicates only the relative probabilities of a set of possible image sources, a particular source (i.e., a particular combination of physical subtense and orientation in the above example) must be selected from the set as the predicted percept. The common way of addressing this issue is to assume that the visual system makes this choice according to the behavioral consequences associated with each perceptual “decision.” The influence of the various consequences is usually expressed in terms of the discrepancy between the decision made and the actual state of the world, which over the full range of the possible choices defines a gain–loss function. Because there is no *a priori* way to model this function (indeed, given the enormous number of variables involved, an accurate gain–loss function for some aspect of vision is extraordinarily difficult to determine), the relative cost of different behavioral responses is assumed. Applying the resulting rule to the posterior probability distribution allows the investigator to state the specific physical properties that, in this conceptual framework, are represented in what observers actually see in response to the stimulus in question. The need for this additional rule in using Bayes’ theorem to predict visual percepts explains why this general approach is referred to as Bayesian decision theory.

Empirical Ranking Theory. The alternative method used in the present analysis is quite different in that the goal is not to make inferences about the properties of the physical world, or indeed any other variable. In this sense, empirical ranking as a means of understanding visual perception does not fit within the framework of probability theory (the larger framework for any Bayesian approach). On the contrary, the underlying conception of vision is that percepts are simply empirically derived constructs that have been successful in guiding behavior, and that they need not (and do not) correspond to any physical property in the world.

This alternative way of exploring how the visual system contends with the problem of stimulus uncertainty is predicated on the idea that observers have evolved to possess a perceptual range for any visual quality determined by the relative frequency of occurrence of all the physical instances pertinent to that quality in past experience. For example, with respect to the perception of angle subtense, the position of any particular angle percept within the full range of angle percepts should accord with how often the

projected subtense has appeared in relation to all other projected angle values that have been experienced by human observers. To predict the percept elicited by a given angle on the retina, the prior distribution, $P(\text{physical subtense, orientation})$, is integrated to generate a marginal distribution (see Fig. 3). Each point in the marginal distribution thus represents the summed probability of occurrence of all the physical conditions that could have generated a specific projected subtense. The cumulative probability distribution derived from this marginal distribution (see Fig. 4A) is, in effect, an empirical scale that orders the full range of projected angle subtense according to past experience. The rank of any given projected angle on this scale is determined by the percentage of all the physical sources that generated projected angles less than the subtense at issue and the percentage that generated greater values. Such ranking indicates how the full range of a retinal image feature is mapped to the full range of the corresponding perceptual space (e.g., the range of apparent angles). The higher the percentage of physical sources that in past experience generated projected angles less than the subtense of the projected angle at issue, the higher that angle ranks on the empirical scale and thus the greater the apparent subtense.

The biological advantage of this strategy of vision is derived from ordering the perceptions of any given visual quality appropriately. Seeing visual qualities according to their empirical rank, whether of geometrical characteristics or other qualities such as brightness or color, is a scheme of vision that maintains in perceptual space the relative similarities and differences among physical objects.

From a more practical perspective, empirical ranking has been used successfully to explain a range of puzzling effects that are difficult to account for in terms of Bayesian decision theory (see, for example, ref. 10). A geometrical example that makes this point is the perceptual effect elicited by the Müller–Lyer stimulus and its variants (see the accompanying article, “The Müller–Lyer Illusion Explained by the Statistics of Image-Source Relationships”). Applying the Bayesian theory in this instance indicates that the average physical length of the sources of the shaft adorned with arrow tails is in fact shorter than the average physical length of the sources of the shaft adorned with arrowheads. A Bayesian approach would therefore predict that a shaft adorned with arrow tails appear shorter than the same shaft adorned with arrowheads, which is the opposite of the actual percept. In contrast, the Müller–Lyer effect and its variants are predicted successfully by empirical ranking.

1. Howe, C. Q. & Purves, D. (2002) *Proc. Natl. Acad. Sci. USA* **99**, 13184–13188.
2. Geisler, W. S. & Kersten, D. (2002) *Nat. Neurosci.* **5**, 508–510.
3. Heyer, D. & Mausfeld, R. (2002) *Perception and the Physical World: Psychological and Philosophical Issues in Perception* (Wiley, New York).
4. Kersten, D. & Yuille, A. (2003) *Curr. Opin. Neurobiol.* **13**, 150–158.

5. Knill, D. C. & Richards, W. (1996) *Perception as Bayesian Inference* (Cambridge Univ. Press, Cambridge, U.K.).
6. Knill, D. C., Kersten, D. & Yuille, A. L. (1996) in *Perception as Bayesian Inference*, eds. Knill, D. C. & Richards, W. (Cambridge Univ. Press, Cambridge, U.K.), pp. 1–21.
7. Maloney, L. T. (2002) in *Perception and the Physical World: Psychological and Philosophical Issues in Perception*, eds. Heyer, D. & Mausfeld, R. (Wiley, New York), pp. 145–189.
8. Mamassian, P., Landy, M. & Maloney, L. T. (2002) in *Probabilistic Models of the Brain: Perception and Neural Function*, eds. Rao, R. P. N., Olshausen, B. A. & Lewicki, M. S. (MIT Press, Cambridge, MA), pp. 13–36.
9. Rao, R. P. N., Olshausen, B. A. & Lewicki, M. S. (2002) *Probabilistic Models of the Brain* (MIT Press, Cambridge, MA).
10. Yang, Z. & Purves, D. (2004) *Proc. Natl. Acad. Sci. USA* **101**, 8745–8750.